What is claimed is:

1. A method of a continuous speech recognition system for discriminatively training hidden Markov models, the method comprising:

performing segmentation and recognition of speech training data using a first set of recognition models so as to form a first model reference state sequence, and a set of first model hypothesis state sequences;

mapping states in the first model reference state sequence to corresponding states in a second set of recognition models so as to form a second model reference state sequence;

mapping states in the set of first model hypothesis sequences to corresponding states in the second set of recognition models so as to form a set of second model hypothesis sequences; and

discriminatively training selected model states in the second set of recognition models using the mapped state sequences.

2. A method according to claim 1, wherein the hypothesis state sequences are represented by a lattice structure.

3. A method according to claim 1, wherein the first set of recognition models are detailed match models, and the second set of recognition models are fast match models.

4. A method of a continuous speech recognition system for discriminatively training hidden Markov models, the method comprising:

for a mixture component of a hidden Markov model state, calculating a gradient adjustment of the standard deviation of the mixture component, and

i. if the calculated gradient adjustment is greater than a first threshold amount, performing an adjustment of the standard deviation of the mixture component using the first threshold, or

ii. if the calculated gradient adjustment is less than a second threshold amount, performing an adjustment of the standard deviation of the mixture component using the second threshold, or else

iii. performing an adjustment of the standard deviation of the mixture component using the calculated gradient adjustment.

5. A method of a continuous speech recognition system for discriminatively training hidden Markov models, the method comprising:

determining correctness of a hypothesized word using pronunciation of the hypothesized word and a corresponding word in a reference text.